

University of Groningen

Modaliteit als parameter: verschillen tussen spontane en geacteerde spraak.

Gilbers, D.G.; Bos, Laura; Heeres, T.; Muller, M.; Wierenga, E.; de Vries, Nienke

Published in:
TABU

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Early version, also known as pre-print

Publication date:
2009

[Link to publication in University of Groningen/UMCG research database](#)

Citation for published version (APA):

Gilbers, D. G., Bos, L., Heeres, T., Muller, M., Wierenga, E., & de Vries, N. (2009). Modaliteit als parameter: verschillen tussen spontane en geacteerde spraak. *TABU*, 38(1-4), 110 - 120.

Copyright

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

Take-down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

Modaliteit als parameter: verschillen tussen spontane en geacteerde spraak

Dicky Gilbers, Laura Bos, Tim Heeres, Menke Muller, Nora de Vries en Ellis Wierenga

Samenvatting

Het onderscheid tussen spontane en geacteerde spraak wordt gekenmerkt door onder andere de positie van pauzes, spreesnelheid en accuraatheid van articulatie. In dit onderzoek tonen wij aan dat er een extra akoestische parameter te onderscheiden valt, namelijk modaliteit. Hieronder verstaan wij meerdere frequentiepieken in een kort spraakfragment. We hebben modaliteit in spraakfragmenten van maximaal 10 seconden gemeten door de toonhoogte per syllabe af te ronden op semitonen en de semitonen te clusteren. Waar de meeste spontane spraak gekenmerkt wordt door één frequentiepiek (de gemiddelde toonhoogte), treffen we in geacteerde spraak meerdere frequentiepieken aan. In een bidirectioneel OT-model van Boersma (2009) zullen we laten zien dat cue constraints een dominante rol spelen in geacteerde spraak maar dat in spontane spraak dat effect ‘overruled’ wordt door lenitieconstraints.¹

1. Inleiding

Luisteraars zijn goed in staat het verschil tussen geacteerde en spontane spraak op te merken (Campbell 2001, 2003; Mathon & de Abreu, 2007). Dit is onder andere te verklaren door een verschil in gebruik van pauzes. Zo blijkt dat pauzes in voorgelezen spraak (of: niet-natuurlijke spraak) korter zijn dan in spontane spraak (of: natuurlijke spraak) (O’Connell & Kowal, 1972; Kowal, O’Connell, O’Brien & Bryant, 1975a). Howel en Kadi-Hanifi (1991) tonen aan dat in voorgelezen spraak minder pauzes voorkomen dan in spontane spraak. Ook de plaats van pauzes is verschillend. In spontane spraak valt 55% van alle pauzes op een grammaticale grens (Henderson, Goldman-Eisler & Skarbek, 1966), waar in voorgelezen spraak de pauzes bijna altijd samenvallen met een grammaticale grens (Levin, Schaffer & Snow, 1982).

¹ We zijn de drie actrices die hebben meegewerkt aan dit onderzoek dank verschuldigd: Mirjam Eringa, Inger Kolman en Lutske Folkerts. Daarnaast willen wij ook graag alle proefpersonen bedanken die hebben meegewerkt aan dit onderzoek.

Levin Schaffer en Snow (1982) lieten 4 onderwijzers verhaaltjes vertellen en voorlezen. Het opgenomen materiaal werd in fragmenten van 10 seconden geknipt. Uit de helft van de fragmenten werd de semantische informatie uit het signaal gefilterd. 11 proefpersonen gaven zowel in de gefilterde als in de niet-gefilterde fragmenten in de meeste gevallen correct aan of het om voorgelezen of vertelde spraak ging. Behalve het verschil in gebruik van pauzes blijkt dat in voorgelezen spraak 1) de spreek snelheid hoger ligt, 2) minder valse starts en herhalingen voorkomen, 3) het ritme constanter is en 4) klinkerverlengingen zeldzamer zijn.

Voorgelezen spraak is niet spontaan en lijkt qua eigenschappen meer op geacteerde spraak. Toch tonen Rusko, Trnka, Darjaa, Kovac en Hamar in 2008 aan dat er ook verschillen zijn tussen voorgelezen en geacteerde spraak. Zij concluderen dat de toonhoogte aan het eind van de zin in het Slowaaks voorgelezen spraak meer omlaag gaat dan in geacteerde spraak. Verder blijkt dat voorgelezen spraak langzamer is dan geacteerde spraak.

Een onderzoek waarin geacteerde en spontane spraak direct met elkaar vergeleken worden is dat van Gunkle (1968). In dit onderzoek zijn 16 mannelijke acteurs geïnterviewd in een opnamestudio. De gesprekken zijn vervolgens uitgeschreven en aan diezelfde acteurs aangeboden om voor te lezen. Ook de voorgelezen spraak is opgenomen. Een groep beoordelaars werd vervolgens gevraagd de fragmenten op een aantal factoren te beoordelen. Gunkle concludeerde dat spontaan klinkende taal langzamer is, tempowisselingen en meer stiltes heeft en minder moeite lijkt te kosten.

Lerdahl en Jackendoff (1983), Gilbers en Schreuder (2002) en Schreuder (2006) tonen aan dat spraak en muziek veel gemeenschappelijke eigenschappen hebben. Muziekstukken met een majeursprong worden waargenomen als vrolijk en muziekstukken met een mineursprong klinken droevig. Schreuder, Van Eerten en Gilbers (2006) publiceerden een onderzoek naar modaliteit in de spraak van de Winnie de Poeh karakters Teigetje (vrolijk) en Iejoor (verdrietig). Zij lieten leerkrachten van een basisschool passages voorlezen uit de verhalen waarin deze twee karakters voorkomen. In de gevallen waarin het bepalen van het type modaliteit mogelijk was, vonden ze bij Teigetje alleen majeur en bij Iejoor alleen mineur.

Luisteraars maken bij het beoordelen van fragmenten op spontaniteit dus gebruik van de volgende eigenschappen in het signaal: spreektempo, aantal en lengte van pauzes en ritme. Wij zouden aan dit rijtje een andere prosodische eigenschap willen toevoegen, te weten: modaliteit. Onder modaliteit verstaan wij meerdere frequentiepieken in het spraaksignaal. Wij zullen in dit onderzoek aantonen dat in geacteerde spraak meer modaliteit voorkomt dan in

spontane spraak. Het type modaliteit hebben we opgesplitst in mineur, majeur en ongedefinieerd en ook apart geanalyseerd.

In de volgende paragrafen zullen we eerst het onderzoek presenteren dat aantoonst dat in geacteerde spraak meer modaliteit voorkomt dan in spontane spraak. In paragraaf 2 wordt de methode toegelicht en in paragraaf 3 worden de resultaten gepresenteerd. In paragraaf 4 volgt een theoretische achtergrond voor het koppelen van onze bevindingen aan Optimality Theory (Prince & Smolensky, 1992). In de discussie in paragraaf 5 wordt hier verder op ingegaan, waarop we afsluiten met de conclusie in paragraaf 6.

2. Methode

Voor het verkrijgen van spontaan vrolijke spraak hebben we bij 22 vrouwen tussen de 18 en 34 korte interviews afgenomen in het Nederlands en die opgeslagen als .wav-bestanden. De geïnterviewden variëren in leeftijd van 18 tot 25 en hebben Nederlands als moedertaal. We hebben ervoor gekozen om alleen vrouwelijke proefpersonen te gebruiken voor dit onderzoek zodat sekse geen factor was in onze verdere analyses.

In het interview vroegen we onder andere naar de leukste film die de geïnterviewde persoon de laatste tijd had gezien en vroegen vervolgens om de leukste scene uit die film te beschrijven. We hebben sommige gesprekken de vrije loop gelaten wanneer we van mening waren dat het spontaan aan de orde gekomen onderwerp ook vrolijke spraak op zou leveren. De geïnterviewden wisten van tevoren dat het interview opgenomen zou worden maar kregen verder geen informatie over het onderzoek. Dit hebben we gedaan om de spontane spraak zo min mogelijk te forceren. Met ‘zo min mogelijk’ bedoelen we dat de spontane spraak af en toe alsnog enigszins geforceerd overkwam omdat sommige geïnterviewden de aanwezigheid van een voice-recorder als nerveusmakend ervoeren. De geïnterviewden zijn na afloop van het interview alsnog geïnformeerd over de inhoud van het onderzoek.

Uit de 22 interviews hebben we naar eigen oordeel 92 vrolijk klinkende fragmenten van 5 tot 10 seconden gehaald. Vervolgens hebben we 8 lijsten samengesteld met 14 vrolijk klinkende fragmenten en 6 fillerfragmenten. Om te zorgen dat de semantische informatie niet zou bijdragen aan het oordeel, hebben we de fragmenten eerst gefilterd zodat de spraak niet meer te verstaan was maar de prosodie — de fundamentele toonhoogte — nog wel duidelijk te onderscheiden bleef. Dit hebben we gedaan met behulp van een FFT-filter in het computerprogramma Adobe Audition 3.0.

We hebben in totaal 116 personen verdeeld over de 8 lijsten gevraagd te beoordelen hoe vrolijk ze de fragmenten van een lijst vonden klinken op een schaal van 1 (helemaal niet vrolijk) tot 5 (erg vrolijk). In de analyses hebben we de grens voor ‘vrolijke spraak’ gezet op 3,4. De 21 fragmenten die als vrolijk werden beoordeeld, hebben we uitgeschreven. Deze 21 fragmenten kwamen van 11 personen variërend in leeftijd van 18 tot 25 jaar.

De uitgeschreven fragmenten hebben we gebruikt om een script met drie verschillende gesprekken samen te stellen, te vinden in Appendix 1. Op deze manier hebben we tot op zekere hoogte de variatie in foneeminventaris van het spraaksignaal en lengte van de fragmenten kunnen controleren. De uitgeschreven fragmenten moesten ‘vrolijk’ worden geacteerd; deze fragmenten hebben we aangevuld met andere stukken tekst die op neutrale wijze moesten worden geacteerd. Hierdoor konden we het script tot een beter passend geheel maken en hoefde de actrice niet alle tekst ‘vrolijk’ te acteren. Door het contrast konden de vrolijke stukken beter worden aangezet. Dit script hebben we vervolgens door 3 verschillende actrices laten acteren, waarvan opnames in .wav-formaat zijn gemaakt. De actrices varieerden in leeftijd en acteerervaring, zie tabel 1. Tijdens de opname hielden de actrices de tekst erbij als geheugensteun.

Tabel 1: Leeftijd en acteerervaring van de drie actrices

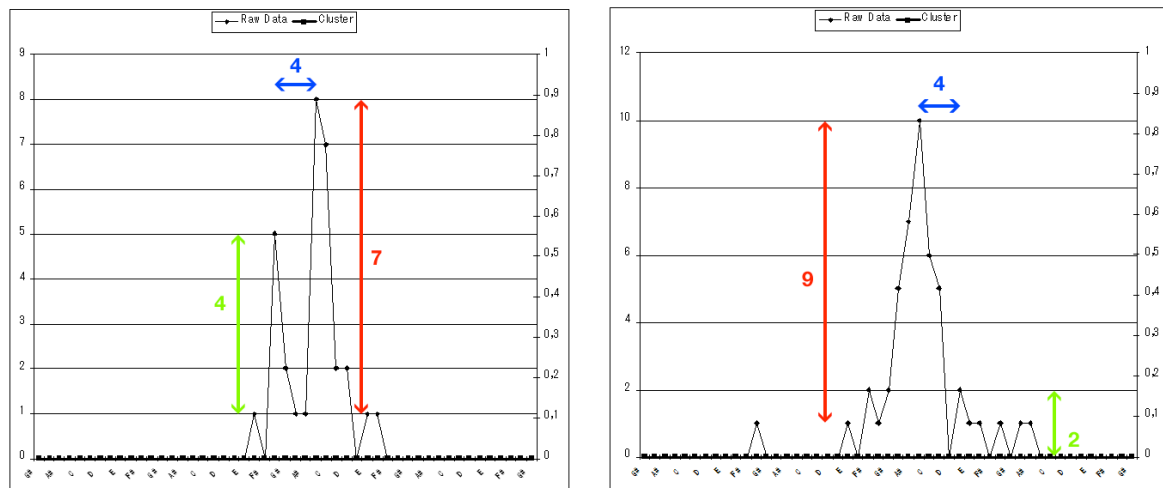
	Leeftijd	Acteerervaring
Mirjam Eringa	33	4 jaar (1 jaar fulltime) van amateur- tot semi-professioneel
Inger Kolman	22	Sinds 14 jaar op verschillende niveaus, tot semi-professioneel
Lutske Folkerts	22	Enkele voordrachten van teksten voor een klein publiek

Op de verkregen opnames hebben we vervolgens dezelfde procedure toegepast als hiervoor beschreven. De 21 fragmenten werden uit de opname geknipt, vervolgens werd er een FFT-filter over de fragmenten heen geplaatst. De gefilterde fragmenten werden verdeeld over 4 lijsten aan totaal 57 personen ter beoordeling aangeboden. De fragmenten die als 3,4 of hoger werden beoordeeld op een schaal van 1 (helemaal niet vrolijk) tot 5 (erg vrolijk) hebben we als ‘vrolijk’ gedefinieerd.

Voor alle als vrolijk beoordeelde fragmenten hebben we met behulp van een script (De Jong & Wempe, 2009) de syllabes gedetecteerd. Dit hebben we gedaan met een script (De Jong & Wempe, 2009) voor het computerprogramma PRAAT (Boersma & Weenink 1992-2005). Er is hierbij een *sampling rate* van 10 milliseconde gebruikt, wat in een vergelijkbaar onderzoek

van Schreuder, Van Eerten en Gilbers (2006) voldoende bleek te zijn. Daarna hebben we met behulp van een script van Cook (Cook, 2002; Cook, Fujisawa & Takami, 2004) de intensiteit en de pitch bepaald. Met een Excel-macro (Cook, 2002; Cook, Fujisawa & Takami, 2004) hebben we vervolgens per fragment de toonhoogtes geclusterd op semitoonafstand. Dit betekent dat de waarden van de toonhoogte omhoog of omlaag werden afgerond naar de dichtstbijzijnde semitoon.² De geclusterde toonhoogtes werden vervolgens door het Excel-macro weergegeven in histogrammen.

We hebben alle histogrammen, zowel die van de spontane als die van de geacteerde spraak, bestudeerd aan de hand van een formule om zo te concluderen of er per fragment sprake was van modaliteit. De formule kan als volgt worden beschreven: er moeten twee pieken te constateren zijn in een histogram dat het resultaat is van de verwerking van de toonhoogtes van de syllables in een bepaald (geacteerd of spontaan) gesproken fragment. De horizontale afstand tussen twee pieken moet minstens 2 semitonen zijn, omdat er anders nooit twee pieken kunnen zijn; de verhouding van de verticale afstanden van beide pieken ten opzichte van het laagst tussen gelegen dal mag niet meer dan 2,5 zijn. Zie figuur 2 voor twee histogrammen waarvan de linker een voorbeeld geeft van modaliteit en de rechter het ontbreken van modaliteit weergeeft.



Figuur 2 : Twee voorbeeldhistogrammen verkregen voor een spraakfragment: links met, rechts zonder modaliteit.

Vervolgens hebben we gekeken naar de eventuele aanwezigheid van majeur- of mineurmodaliteit. We gaan er van uit dat de hoogste piek de grondtoon is. Voor majeur is de

² Deze abstractie is vergelijkbaar met die van fonologen wanneer zij verschillende allofonen omschrijven als realisaties van hetzelfde foneem

afstand tussen twee toonhoogtepieken 4 semitonen terwijl dit voor mineur 3 is. Bij een afstand groter dan 4 semitonen is de modaliteit ongedefinieerd.

3. Resultaten

Van de 92 geselecteerde gefilterde fragmenten met spontane spraak zijn er 21 als 3,4 of hoger beoordeeld op een schaal van 1 (helemaal niet vrolijk) tot 5 (heel vrolijk). Van de 63 gefilterde fragmenten met geacteerde taal (21 per actrice) zijn er 20 beoordeeld als 3,4 of hoger op deze schaal. Van de actrice Mirjam waren dit er 6, van Inger waren dit er 3 en van Lutske waren dit er 11.

In geacteerd vrolijke spraak zit meer modaliteit (75%) dan in spontaan vrolijke spraak (43%; $p < .05$; Fisher's Exacte Toets ééNZijdig). Van de 20 fragmenten geacteerd vrolijke spraak waren er 5 zonder modaliteit, 10 met ongedefinieerde modaliteit, 4 met majeur modaliteit en 1 met mineur modaliteit. De meeste gevonden modaliteit is dus ongedefinieerd.

Een Kruskal Wallis test laat geen significant verschil zien tussen de vrolijkheidsoordelen voor de 3 types modaliteit (majeur, mineur en ongedefinieerd; $\chi^2=2,5$; $df = 2$, $p=0,3$) bij geacteerd vrolijke spraak.

4. Theoretische achtergrond

Iedere taal, en daarmee de bijbehorende grammatica, is anders. In 1993 introduceerden Prince en Smolensky een Optimality Theory (OT) waarmee dit kan worden uitgelegd. In OT wordt er vanuit gegaan dat er een verzameling universele taalregels bestaat (Gilbers & de Hoop, 1998), de *constraints*. Dit zijn beperkingen die aan grammatica's worden opgelegd, en deze constraints functioneren in een bepaalde rangschikking. Deze rangschikking verschilt per taal waardoor iedere grammatica er anders uitziet. Anders gezegd: iedere taal kent een andere waarde toe aan de verschillende beperkingen, waardoor sommige constraints in de ene taal een grotere prioriteit hebben en dus hoger in de rangorde komen te staan dan in de andere taal. Constraints kunnen geschonden worden en met elkaar wedijveren.

Een voorbeeld van een constraint is de beperking NoCoda, die stelt dat een woord geen consonant als laatste klank mag hebben. Het woord *boom* schendt die beperking dus, tegenover het woord *video* dat die constraint niet schendt. Ook een verschil in klanken maakt onderdeel uit van de constraintrangschikking. Denk bijvoorbeeld aan de verschillende

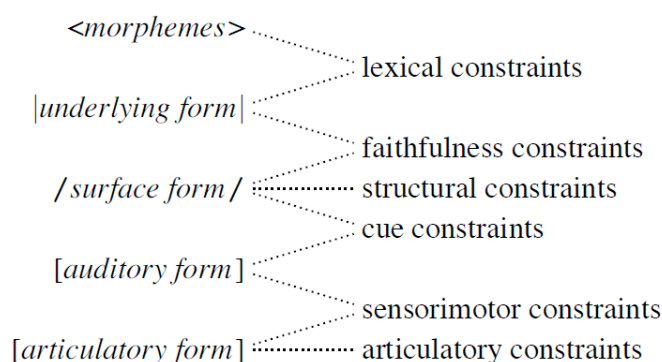
manieren waarop een *i* uitgesproken kan worden. Al deze verschillende manieren worden kandidaten genoemd.

Belangrijke constraints voor ons onderzoek zijn *faithfulness* constraints, *cue* constraints (minimal distance) en *articulatory* constraints. Faithfulness constraints verbieden de modificatie van input-representaties (Hale & Reiss 1996), oftewel: deze constraints vereisen dat de oppervlakte vorm gelijk is aan de onderliggende vorm.

Cues zijn ideaalpatronen die gekoppeld zijn aan een fonologische representatie van een klank en zowel bij productie als perceptie gebruikt worden. Hierbij moet gedacht worden aan bijvoorbeeld de hoogte van een formant of de tijdsduur van een klank. De beslissing dan van de spreker of hoorder om een bepaalde klank te interpreteren wordt bepaald door de aanwezige ontwikkelde kennis van constraints. Cue constraints zijn niet binair, maar moeten uit bijvoorbeeld het auditieve spectrum gehaald worden.

Articulatory constraints zijn voorwaarden die alleen worden gebruikt in gesproken taal, de naam zegt het al: het zijn voorwaarden die pleiten tegen articulatory effort, en wedijveren daardoor mogelijkwerwijs met de faithfulness constraints.

Het bidirectionele OT-model (Boersma, 2005) bestaat uit verschillende categorieën constraints die elkaar in een bepaalde volgorde opvolgen tijdens het perceptieproces dan wel productieproces van taal. Dit is te zien in het model in figuur 2.



Figuur 2: Het OT-model (Boersma, 2005)

Het model gaat, zoals gezegd, twee richtingen op: die van productie en die van perceptie. Daarom wordt het ‘bidirectioneel’ genoemd. De productie wordt voorafgegaan door de perceptie, maar hoeft uiteindelijk niet dezelfde waarde te hebben als de perceptie, omdat er constraints zijn die met elkaar wedijveren op het hele traject van productie.

Bij perceptie moet de luisteraar allereerst continue, auditieve informatie met behulp van cue constraints omzetten in een meer abstracte, mentale representatie: de discrete fonologische oppervlaktevorm (SF). Deze SF wordt geëvalueerd aan de hand van structural constraints. De luisteraar moet deze SF via faithfulness constraints mappen met de onderliggende vorm (UF) die geëvalueerd wordt aan de hand van lexical constraints. Op deze wijze wordt betekenis gegeven aan de fonologische vorm.

Productie bestaat uit de mapping van UF met ArtF zonder tussenstops op de tussenliggende niveaus. Evaluatie geschiedt dan aan de hand van faithfulness, cue, sensorimotor en articulatory constraints (Boersma, 2005).

Modaliteit als cue

Dit onderzoek heeft als doel aan te tonen dat modaliteit ook tot de cues hoort. Het is een cue met een graduele schaal, wat inhoudt dat er geen specifiek omslagpunt is van niet-modaal naar modaal. Voor het bepalen van modaliteit is de volgende formule opgesteld: er moeten twee pieken te constateren zijn in een histogram dat het resultaat is van de verwerking van de toonhoogtes van de syllables in een bepaald (geacteerd of spontaan) gesproken fragment (zie voor het precieze proces van het omzetten van gesproken taal naar histogrammen paragraaf 2: Methode). De horizontale afstand tussen twee pieken moet minstens 2 semitonen zijn, omdat er anders nooit twee pieken kunnen zijn; de verhouding van de verticale afstanden van beide pieken ten opzichte van het laagst tussen gelegen dal mag niet meer dan 2,5 zijn. Deze formule is enigszins arbitrair, waarop nog verder ingegaan zal worden in de discussie aan het eind van dit artikel. Voorwaarde is in ieder geval dat er sprake is van twee toonhoogtepieken.

Een voorbeeld van een ander gradueel verloop van cues vinden we bij Boersma (2004). In zijn onderzoek zijn twee groepen native speakers van het Spaans onderzocht. De ene groep leerde Schots en de andere ‘Southern English’. Het ging hier met name om de /e/ en /i/ klanken en het aanleren hiervan. De grens tussen beide klanken ligt in het Spaans op een andere plek dan in de beide doeltalen. Hierbij spelen twee cues een belangrijke rol, namelijk *spectrum* en *duration*, waarbij de eerstgenoemde uitgedrukt wordt in Hertz (c.f. frequentie) en *duration* staat voor tijdsduur. Er is een gradueel verloop tussen deze beide cues. Deze loopt van *exclusively spectrum* tot *exclusively duration* en daartussenin bevinden zich onder andere cues als *mainly spectrum* en *mainly duration*. Op basis van deze twee cues delen native speakers van het Schots en het Southern English de /e/ en /i/ klanken in; de native speakers van het Spaans doen hetzelfde bij deze klanken in hun eigen taal. Als vervolgens deze Spanjaarden dit onderscheid moeten maken bij één van de beide genoemde dialecten van het

Engels, moeten ze hun cue constraints bijstellen, omdat die zoals gezegd verschillen tussen beide talen. Dit gebeurt volgens het in het begin van deze paragraaf besproken *gradual learning algorithm*.

In het onderzoek van Boersma (2004) komt naar voren dat cue constraints aangeleerd zijn; de verschillen tussen talen onderling leveren hier ook bewijs voor. In het Japans en het Engels vinden we doorgaans in de histogrammen drie toonhoogtepieken tegenover twee pieken in het Nederlands; ook dit onderling verschil levert dus bewijs dat het gaat om een aangeleerde cue.

Uit het voorgaande kan worden geconcludeerd dat modaliteit de default is, wat wil zeggen dat de maximale vorm van de modaliteit de modaliteit is die opgeslagen ligt in de templatens in het brein van de mens. Modaliteit komt in spontaan vrolijke spraak in afgezwakte vorm of zelfs geheel niet voor doordat deze cue geschonden wordt door andere constraints. Onze hypothese luidt dan ook dat in spontaan vrolijke spraak significant minder modaliteit voorkomt dan in geacteerd vrolijke spraak. Er is gekozen voor geacteerd vrolijke spraak, omdat juist hierbij gebruik wordt gemaakt van de templatens. Dat er in spontane spraak afgeweken wordt van die default, valt te wijten aan schendingen door andere, voornamelijk articulatorische constraints.

Constraints en modaliteit

Naast de aan het begin van deze paragraaf besproken cue constraints kent het bidirectionele model (Boersma, 2005) nog twee andere constrainttypen die bepalend kunnen zijn voor de uiteindelijke perceptie van geacteerd vrolijke en spontaan vrolijke spraak. Eerst zullen we een verklaring geven voor het voorkomen van deze constraints bij dit onderzoek en vervolgens zullen we de constraints in een model toepassen om zo een beschrijvende ranking te verkrijgen.

Articulatory effort

Articulatorische lenitieprocessen vanuit een OT perspectief zijn beschreven door onder andere Kirchner (2001) en Flemming (2004). Hierbij moet gedacht worden aan *voicing*, *degeminatie*, *elisie*, *pallatalisatie*, *nasaalassimilatie* etc. Het gebruik van deze processen stelt de spreker in staat om met zo weinig mogelijk biomechanische energie toch verstaanbaar te zijn. Dit fonetisch imperatief zorgt voor lenitie oftewel de minimalisatie van *effort-cost* bij een articulatorische beweging. Alle processen die hierdoor veroorzaakt worden, kunnen in één

constraint gevat worden, namelijk LAZY (Kirchner, 2001). Al deze processen wedijveren met constraints die LAZY blokkeren.

We willen LAZY als concurrent aanvoeren van cue constraints, omdat het mogelijk is om met een verminderde spierspanning, die ook verantwoordelijk is voor andere lenitieprocessen, zich toch voor de luisteraar voldoende begrijpelijk uit te drukken. Voor ons onderzoek zou dan gelden dat een luisteraar met een ongeschonden LAZY constraint toch spraak als vrolijk zou kunnen beoordelen.

Preserveness

LAZY is in de vorm zoals hierboven beschreven nog onbeperkt en zou mogelijkerwijs met een meest minimale effort een uiting zonder vrolijkheid kunnen worden, terwijl deze uiting wel als zodanig bedoeld is. Er is een andere constraint nodig om de tot nu toe infiniete LAZY constraint zodanig te begrenzen dat er geen misverstand kan bestaan over de intentie van de uiting. Deze uitingsdrempel kan gerealiseerd worden door de spreker indien hij zich houdt aan de getrouwheid van zijn uiting. Dit houdt in dat zijn vrolijke uiting zoals bedoeld in de input tot realisatie komt in de output: de uiteindelijke spraak. Indien deze PRESERVENESS(happiness) constraint geschonden wordt zal LAZY de uiting domineren en is deze niet meer als vrolijk kwalificeerbaar. Deze drempel sluit aan bij het coöperatieve principe van Grice (ref.), waarvoor geldt dat de spreker met minimale effort toch de maximale duiding bij de toehoorder tracht te bereiken.

Waar de uitingsdrempel precies ligt kan per gesprek verschillen. Moedertaalsprekers zullen in een spontane conversatie met elkaar minder moeite doen om toch begrijpelijk over te komen dan wanneer een moedertaalspreker converseert met een spreker die de gesproken taal niet vloeiend verstaat. In het laatste geval zal de spreker zich moeten aanpassen wat betreft de mate van ongeremdheid van lenitieprocessen en zal bijvoorbeeld beter moeten articuleren.

Cue constraint voor modaliteit

Om geen onderlinge verschillen te laten bestaan bij luisteraars over de vrolijke intentie van een uiting van een acteur probeert deze de meest onderscheidende waarden (Flemming, 2004) binnen een uiting te gebruiken. Dit ideaalpatroon kent tonale eigenschappen die gebruikt kunnen worden om een cue constraint voor modaliteit te creëren.

De geponeerde formule van modaliteit, zoals beschreven in paragraaf 2: Methode, kan op de volgende manier tot een concrete constraint geformuleerd worden:

Modaliteitsafstand: Een minimale horizontale afstand van twee semitonen tussen de hoogste toonpieken en een minimale verticale verhouding van deze toonpieken is niet groter dan een factor 2.5.

Slechts als deze beide variabelen aanwezig zijn in een uiting kan gesproken worden van een ongeschonden cue constraint *modaliteitsafstand*. Om de range van de constraint beter te beperken indachtig de formulering van Boersma (2008) zal deze in het model negatief geformuleerd zijn met als variabelen de horizontale en verticale verhoudingen. Modaliteit wordt zo geformuleerd: *MODDISTANCE (<2.sem.hor, >2,5.vert).

Realisatie OT-model

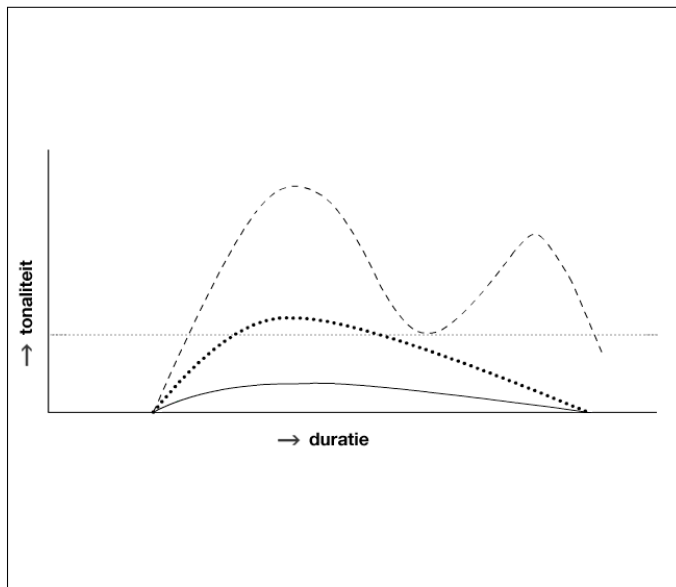
Nu de concurrerende constraints geformuleerd zijn, kunnen zij gerangschikt worden voor elke kandidaat van vrolijke taal. Dit zijn de spontane spreker en de acterende spreker. Een derde spreker kan theoretisch ook toegevoegd worden, namelijk een spreker die zich vrolijk tracht te uiten, maar LAZY ongeschonden laat en daardoor niet voldoet aan PRESERVENESS(happiness). Deze spreker heeft zo weinig tonaal verschil dat hij als monotoon gekwalificeerd kan worden door een luisteraar.

Voor de spontane spreker geldt dat hij zoals gesteld met zo weinig mogelijk moeite vrolijk wil klinken. Dit houdt voor het model in dat de MODDISTANCE constraint geschonden wordt door de LAZY constraint, die zelf ongeschonden blijft, evenals de PRESERVENESS (happiness) constraint om zo de vrolijke intentie te behouden. Door de hogere ranking van de laatstgenoemde constraint is het dus niet meer mogelijk voor de LAZY constraint om zo sterk te werken dat de uiting niet meer voldoet aan de vrolijkheid. Er is dus automatisch al een restrictie op de LAZY constraint, oftewel de ondergrens is omhooggetrokken door de PRESERVENESS (happiness) constraint.

Voor de acterende spreker geldt dat alleen LAZY geschonden wordt, daar de spieren extra worden aangespannen voor het gewilde effect. MODDISTANCE en PRESERVENESS (happiness) worden niet geschonden. Deze laatste is redundant in geacteerde spraak. Het valt immers niet te verwachten dat een uiting met modaliteit de PRESERVENESS (happiness) constraint zal schenden, want cue constraints maximaliseren de *effort cost* en dus is er per definitie geen schending van PRESERVENESS (happiness) bij geacteerde spraak.

Tenslotte geldt voor de derde, spontane spreker dat MODDISTANCE en PRESERVENESS (happiness) geschonden worden, omdat de input niet hetzelfde is als de

output. Deze spreker uit zich monotoon met niet herkenbare vrolijke intenties. Het spectrum waarbinnen deze drie kandidaten interacteren zou er als volgt uit kunnen zien:



De gestreepte curve representeert een uiting met modaliteit, die net als de gestippelde, spontaan vrolijke curve, boven de grijze stippellijn, PRESERVENESS(happ.) uitkomt. De solide lijn representeert een monotoon vrolijke spreker, die niet voldoet aan de PRESERVENESS(happ.) drempel en dus niet als vrolijk overkomt.

OT-model

Om de precieze ranking en schendingen aan te kunnen geven maken we gebruik van twee OT-tabellen, één voor een spontane uiting en één voor een geacteerde uiting die beiden in hun respectievelijke omgeving vrolijk dienen te klinken. In het model worden de kandidaten getoetst aan de spontaan vrolijke en geacteerd vrolijke uiting zoals die geïnterpreteerd kan worden door de luisteraar.

(1) rangschikking en schending van spontaan vrolijke uiting

<i>Vrolijke taal</i>	*MODDIST (<2.sem.hor, >2,5.vert)	LAZY	FAITH
→Geacteerd vrolijk		*	
Spontaan vrolijk	*!		
Spontaan niet-vrolijk	*!		*

Als de vrolijke uiting verstaan wordt bij een CUE die LAZY domineert, dan is dit voor de luisteraar het signaal dat de vrolijke uiting geacteerd is.

(2) rangschikking en schending van geacteerd vrolijke uiting

<i>Vrolijke taal</i>	FAITH	LAZY	*MODDIST <2.sem.hor, >2,5.vert
→Spontaan vrolijk			*
Geacteerd vrolijk		*!	
Spontaan niet-vrolijk	*!		*

Deze rangschikkingen hangen samen met het aangeleerde verwachtingspatroon van de luisteraar. Indien een spreker in een spontane omgeving zich niet zou houden aan de rangschikking van de eerste tabel zou de luisteraar dit kunnen opvatten als een niet oprechte uiting, maar als gespeeld.

Samengevat kunnen we zeggen dat drie constraint-types kunnen interacteren in twee verschillende omgevingen, spontaan en geacteerd. Het ideaalpatroon modaliteit wordt alleen aangesproken bij geacteerde taal, maar schendt daarbij lenitieprocessen. De afweging die een spreker maakt is een strijd tussen het zo onderscheidend mogelijk op luisteraars overkomen wat veel energie kost, of juist met verminderde energie net onderscheidend genoeg overkomen zodat de luisteraar nog in staat is de uiting te begrijpen.

5. Discussie

In deze paragraaf zullen we de gevonden resultaten van ons onderzoek bediscussiëren en betrekken op het bidirectionele OT-model (Boersma, 2005) dat we in de vorige paragraaf hebben toegelicht.

Zoals we verwachtten bevat geacteerd vrolijke spraak meer modaliteit dan *spontaan* vrolijke spraak. Het bewijzen van deze belangrijkste hypothese van dit onderzoek impliceert dat er een ideaalpatroon voor vrolijkheid bestaat dat in het OT model als de nieuwe cue constraint MODDISTANCE wordt gerepresenteerd en die afhankelijk van de omgeving en het verwachtingspatroon van luisteraars hoger of lager gerangschikt is dan de articulatorische lenitieconstraint LAZY. Net als andere eigenschappen zoals een versneld spreektempo, een verminderd aantal pauzes en een verminderde lengte van pauzes is 'modaliteit' een extra parameter die geacteerd vrolijke spraak definieert.

In een tweede analyse hebben we het type modaliteit geanalyseerd en een tendens gevonden die we niet met statistiek konden analyseren. Hiervoor waren te weinig fragmenten beschikbaar. De meeste fragmenten hadden een ongedefinieerde modaliteit en slechts één fragment had mineur modaliteit. Ook om te bepalen of geacteerd vrolijke spraak met majeur modaliteit als vrolijker beoordeeld wordt dan met mineur of ongedefinieerde modaliteit zijn er meer fragmenten nodig.

In het onderzoek van Schreuder, Van Eerten en Gilbers (2006) werd er voorgelezen en is er geen enkel fragment van de vrolijke Teigetje gevonden waarin mineur modaliteit zat. Zij hebben ongedefinieerde modaliteit buiten beschouwing gelaten, omdat dit buiten het doel van hun onderzoek viel. Het enige tegenvoorbeeld dat we hebben gevonden kan op louter toeval berusten: vervolgonderzoek zal dit moeten aantonen.

Gunkle (1968) heeft in zijn onderzoek de fragmenten spontane spraak en geacteerde spraak (voorgelezen en gememoriseerd) van steeds dezelfde persoon gebruikt. Hij onderzocht minder gericht de verschillen tussen spontane en geacteerde spraak. Wij hebben met dit onderzoek aangetoond dat er een duidelijk verschil is tussen spontane en geacteerde spraak. In ons onderzoek is gebruik gemaakt van twee verschillende groepen mensen voor het produceren van de spontane en geacteerde taal, omdat er met onze opzet veel meer spontane taal dan geacteerde taal beschikbaar moest zijn.

De formule waarvan in dit onderzoek gebruik is gemaakt is enigszins arbitrair. Verder onderzoek zou moeten uitwijzen of de formule nog enige nuancering danwel anderszins aanpassing behoeft. Wij zijn van mening dat de formule zoals deze door ons is opgesteld, zeker een toevoeging is in het onderzoek naar het vaststellen van modaliteit. Het grootste voordeel is, dat de formule consequent is toe te passen. De horizontale afstand is goed te beredeneren, want bij een afstand van 1 semitoon zou er nooit sprake kunnen zijn van twee pieken, maar slechts een schuine lijn. De door ons gestelde factor van de verticale afstand is dus datgene wat de formule arbitrair maakt, maar de reden voor het feit dat wij gekozen hebben voor een vergelijking tussen hoogste en de op één na hoogste piek ten opzichte van het tussen gelegen laagste punt is vanwege gevallen waarbij de op één na hoogste piek niet veel hoger ligt dan het dal en dus in onze ogen geen verschil laat zien ten opzichte van de frequentie van de omliggende toonhoogtes in het diagram en daarbij door toeval te verklaren zou zijn. Op deze manier zit er alleen modaliteit in een fragment als de histogrammen een duidelijke afstand laten zien tussen twee duidelijke pieken en een grondtoon die niet door toeval ontstaan kunnen zijn.

6. Conclusie

Naar aanleiding van de uitkomsten van ons onderzoek zijn wij van mening dat er een constraint met betrekking tot modaliteit bestaat. Dit stellen wij naar aanleiding van het onze bevinding dat geacteerd vrolijke spraak meer modaliteit bevat dan *spontaan* vrolijke spraak. We stellen daarom ook voor om modaliteit toe te voegen aan de lijst met kenmerken van geacteerde spraak.

Literatuur

- Boersma, P. (2005). Prototypicality judgments as inverted perception. In: Fanselow, G. (2006), *Gradience in Grammar: generative perspectives*. Oxford University Press, United Kingdom, 167-170.
- Boersma, P. (2006). A programme for bidirectional phonology and phonetics and their acquisition and evolution. Earlier version: Handout LOT Summerschool and Jadertina Summerschool. *Rutgers Optimality Archive* 868, 2006/09/12.
- Boersma, P. & Escudero, P. (2008). Learning to perceive a smaller L2 vowel inventory: an Optimality Theory account. In: In Peter Avery, Elan Dresher & Keren Rice (eds.): *Contrast in phonology: theory, perception, acquisition*, 271-301. Berlin: Mouton de Gruyter.
- Boersma, P. & D. Weenink (1992-2004). Praat: a system for doing phonetics.
<<http://www.praat.org>.>
- Campbell, N. (2001). The recording of emotional speech (JST/CREST database research) In: *Proceedings frog COCOSDA Workshop*. Taejon, Korea.
- Cook, N.D. (2002). *Tone of voice and mind. The connections between intonation, emotion, cognition and consciousness*. Amsterdam: John Benjamins.
- Cook, N.D., T.Fujisawa & K. Takami (2004). Application of a psycho-acoustical model of harmony to speech prosody. *Proceedings of speech prosody*. Nara, Japan, 147-150.
- Eerten, L. van, Schreuder, M. & D. Gilbers (2006). Mineur en Majeur in Emotionele Spraak [Minor and Major in Emotional Speech]. *TABU* 35 (1/2), 1-15.
- Escudero, P. & Boersma, P. (2004). Bridging the gap between L2 speech perception research and phonological theory. *Studies in Second Language Acquisition* 26: 551-585, Cambridge University Press.

- Gilbers, D. G., Bos, L. S., Heeres, T., Muller, M., Vries, N. de & Wierenga, E. (2010). Modaliteit als parameter: verschillen tussen spontane en geacteerde spraak. *TABU*, 38, p. 110-120.
- Flemming, E. (2006). *The role of distinctiveness constraints in phonology*. Unpublished manuscript, MIT.
- Gunkle, G. (1968). An experimental study of some vocal characteristics of spontaneity in acting. *Communication Monographs*, 35 (2), 159 – 165.
- Gilbers, D. en Hoop, H. de (1998). Conflicting constraints: an introduction to Optimality Theory. *Lingua* 104, 1-12.
- Gilbers, D. & M. Schreuder (2002). *Language and music in optimality theory*. [ROA 571 0103].
- Grice, Paul (1975). "Logic and conversation". In *Syntax and Semantics, 3: Speech Acts*, ed. P. Cole & J. Morgan. New York: Academic Press. Reprinted in *Studies in the Way of Words*, ed. H. P. Grice, pp. 22–40. Cambridge, MA: Harvard University Press (1989)
- Gunkle, G. (1968). An experimental study of some vocal characteristics of spontaneity in acting. *Communication Monographs*, 35 (2), 159 – 165.
- Hale, M. & C. Reiss (1996). *The Initial Ranking of Faithfulness Constraints in UG*. Concordia University, Montréal, p. 2.
- Henderson, A., Goldman-Eisler, F. & Skarbek, A. (1966) Sequential temporal patterns in spontaneous speech. *Language and Speech*, 9, 207-216.
- Howell, P. & Kadi-Hanfini, K. (1991). Comparison of prosodic properties between read and spontaneous speech material. *Speech communication*, 10, 163-169.
- Jong, N.H. De & Wempe, T. (2009). PRAAT script to detect syllable nuclei and measure speech rate automatically. *Behavior Research Methods* 41 (2), 385 - 390.
- Kirchner, R. (2001). *Phonological contrast and articulatory effort*. In: L. Lombardi (ed.), *Segmental Processes in Optimality Theory*, Cambridge University Press, 79-117
- Kowal, S., O'Connell, D.C., O'Brien, E.A. & Bryant, E.T. (1975a). Temporal aspects of reading aloud and speaking: Three experiments. *American Journal of Psychology*. 88, 549-569.
- Levin, H., Schaffer, C.A. & Snow, C. (1982). The prosodic and paralinguistic features of reading and telling stories. *Language and Speech*, 25, 43-54.
- Lerdahl, F. & R. Jackendoff (1983). *A generative theory of tonal music*. The MIT Press, Cambridge, Massachusetts, London, England.
- Mathon, C. & Abreu, de S. (2007). Emotion from speakers to listeners: perception and prosodic characterization of affective speech. In, Müller (ed.): *Speaker Classification*, LNAI, 70-82.

- Gilbers, D. G., Bos, L. S., Heeres, T., Muller, M., Vries, N. de & Wierenga, E. (2010). Modaliteit als parameter: verschillen tussen spontane en geacteerde spraak. *TABU*, 38, p. 110-120.
- McCarthy, John J. (2003). OT constraints are categorical*. In *Phonology*, Cambridge University Press, United Kingdom.
- O'Connell, D.C. & Kowal, S. (1972). Cross-linguistic pause and rate phenomena in adults and adolescents. *Journal of Psycholinguistics Research*, 1, 155-164.
- Prince, A. en Smolensky, P. (1993). *Optimality Theory. Constraint Interaction in Generative Grammar*. ROA Version, 2002.
- Rusko, M., Trnka, M., Darjaa, S., Kovac, R. & Hamar, J. (2008). Modelling acoustic parameters of prosody for read and acted-speech synthesis. *Acoustics, Paris 2008*.
- Schreuder, M. (2006). *Prosodic processes in speech and music*. PhD dissertation. University of Groningen.
- Tesar, B. & P. Smolensky (1998), *Learnability in Optimality Theory*, 236-246.

Appendix 1

Gesprek 1

B: Hoe bevalt je nieuwe baantje? Is het beter dan dat schoonmaken?

A: (*Schertsend/ vrolijk*) Baal ik wel een beetje van, want daar verdien ik 10 euro en hier verdien ik 5 euro, dus dat is veel...ja, ik heb de...ik ben 21 en dan is mijn minimumloon 5 euro en 3 cent.

Maar dat vorige baantje nam ik dan ook niet zo serieus. Ik heb een keer eh (*vrolijk*) al die kussentjes van de bank geplukt en ik lag echt...hier voor mijn voeten een kussen, twee...twee of vier van eh die kussentjes voor eh voor m'n lichaam en dan...twee kussens voor m'n hoofd.

B: Maar nu loop je dus stage voor je journalistiekmaster, samen met Carola.

A: (Neutraal) Ja dat is stukken beter dan dat schoonmaakbaantje. Laatst zouden we bijvoorbeeld verslag gaan doen vanaf de landelijke scoutingdag. We hadden toen nog opname-apparatuur nodig.

(*vrolijk*) Carola echt zo... op eeh internet zei van ehm: wie heeft zo'n ding? En iemand echt zo: 'Ja gebruik gewoon je MP3'. Zit je daar bij een interview haha. Kun je echt niet maken.

B: Heb je ook college gehad over het geven van interviews?

A: (*vrolijk*) 't Was altijd een beetje saai hahaha. Die doce-eeh die lerares kon het niet zo goed brengen.

B: Dus hoe ging dat dan toen je op de scoutingdag was?

A: We besloten om ons voor te doen als scouts zodat we niet zo zouden opvallen als journalisten. We kwamen terecht in de kookploeg en daar hebben we echt een mooie dag beleefd. Lekker gekookt ook zeg! Ravioli gemaakt, met ehh

(*vrolijk*) Spinazie, en dan eh dus eh zo erin doen, en dan effe drie minuutjes koken en dat was eh, very nice yes! Parmezaanse kaas..

Gilbers, D. G., Bos, L. S., Heeres, T., Muller, M., Vries, N. de & Wierenga, E. (2010). Modaliteit als parameter: verschillen tussen spontane en geacteerde spraak. *TABU*, 38, p. 110-120.

(neutraal) Nou goed ik weet niet wat er verder in moest, doet er niet toe. Ohja, vergeet ik bijna:

(Vrolijk): en Maxima kwam langs, terwijl ik stond te kneden. En we dachten nog even: ze komt hierheen! Ze komt hierheen! Maar ze liep gewoon voorbij. Dus dat was wel grappig want er stond overal van tevoren van ja Maxima komt langs, doe vooral, ga gewoon door waar je mee bezig bent! Hahaha terwijl echt zo'n meute mensen klik-klik-klik-klik-klik-klik-klik-klik-klik haha. Dat kan je ook niet vragen van mensen, dat doen ze toch niet.

Ja ik ben gewoon doorgegaan met het kneden van het deeg hoor. Ik dacht wooh, hai.

(Neutraal) Ik vind het koningshuis niet zo boeiend.

Gesprek 2

A: (vrolijk) Het was weer leuk deze zomer in Pieterburen. Al die zeehondjes en nou, ik ben eeh deze vakantie verliefd geworden... op Thijs! En dat is nog steeds, dat is vet leuk!.

A: (neutraal) Ja weet je, hij bleek ook bij die dansworkshops te zijn. Het leek me in eerste instantie niks en *(vrolijk)* ik dacht ook van dat robotachtige krijgen we, maar dat was helemaal niet zo. Maar was wel echt superleuk en we hebben houseworkshops gehad...was wel leuk hoor, ja ik eh ik wist ook niet wat ik daar van moest denken, ik had me ook niet ingeschreven,

A: (neutraal) want je moest wel betalen voor die workshops. En dat was op zich naar verhouding helemaal niet zo duur,

A: (vrolijk) Maar toen waren er uiteindelijk bij de houseworkshops zo weinig inschrijvingen dat die ene jongen zei, nou ik doe het wel gratis...toen kwam echt iedereen bij, iedereen meedoen, vet sneu eigenlijk, beetje lullig, maar goed. En bij de workshop buiten heb Thijs voor het eerst gezien en hij stond hier en wij stonden daar en hij huppelde zo langs en huppelde hier zo voor en alles huppelde mee...hij had één huppelmoment en hij maakte precies de bocht en in die bocht zaten wij..

B: (vrolijk) Wat leuk! Moet wel een hilarisch tafereel opgeleverd hebben!

A: (vrolijk) Jaaa Inderdaad! Als ie dat dansje gaat doen... geweldig! kheb zo hard gelachen de eerste keer dat ik 'm zag.. Ook best knap van de leraar, gewoon, nou dat je een hele voorstelling neerzet eh voor mensen en eeh, gewoon het eindresultaat.

Gesprek 3

A: (Neutraal) Het is hier best saai in Groningen.

B: (Neutraal) Ja dat kan ik me voorstellen, heb ook al weken niets te doen.

A: (*vrolijk*) Ik verveel me.. haha dus nou ga ik naar Thesinge, dat is ook nog sportief fietsen haha. Daar is vanavond een groot feest. Ik ben daar al een keer eerder geweest. Supermooi gewoon.

B: (*vrolijk*) Vertel vertel, ik heb daar al veel supermooie verhalen over gehoord

A: (*vrolijk*) Ja en dan echt zie je echt zo hiiee.. Sta je daar met z'n allen weet je wel, van dit is mooi! Maar volgens mij wordt het echt een leuke borrel, want er komen echt vet veel mensen. De vorige keer dat ik daar was, ongeveer een jaar geleden, deden we een spel en mijn vader moest op een gegeven moment...ik weet niet wat hij moest tekenen, storm ofzo, of regen. Hij had een heel blaadje volgekliederd met windslierten. Mijn pa was niet te te houden trouwens, want toen ging hij verder enneh... echt PRECIES bij die volgende stop stonden die twee kerels met dat ehhe. Nou dat was zeg maar een standje. Daar moesten we allemaal meespelen. Er stond een aantal plasticflessen en toen gingen we met een watermeloen ze omgooien. We hebben zelden zo hard gelachen en ik heb direct gebeld met mijn moeder, want dacht ik, nou, dit moeten ze weten, dit moeten ze weten hoe leuk dit is... enne...toen hebben we dat spel gedaan...en nu wil mijn moeder het voor haar verjaardag hebben.